

(19)日本国特許庁(J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平5-11799

(43)公開日 平成5年(1993)1月22日

(51)Int.Cl.⁵

G 1 0 L 9/00
9/18

識別記号

庁内整理番号
J 8946-5H
E 8946-5H

F I

技術表示箇所

審査請求 未請求 請求項の数 4(全 13 頁)

(21)出願番号 特願平3-167270

(22)出願日 平成3年(1991)7月8日

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中1015番地

(72)発明者 加藤 雅子

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(72)発明者 田中 良紀

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(72)発明者 谷口 智彦

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(74)代理人 弁理士 大宮 義之 (外1名)

最終頁に続く

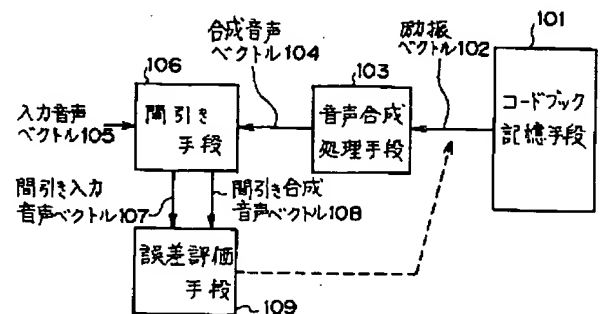
(54)【発明の名称】 音声符号化方式

(57)【要約】

【目的】 本発明は、合成音声の品質の劣化を最小限に抑えつつ誤差評価時の演算量を削減することにより音声符号化処理の高速化を図ることを目的とする。

【構成】 コードブック記憶手段101からは励振ベクトル102が選択される。音声合成処理手段103は、励振ベクトル102を入力として線形予測分析に基づく音声合成処理を行う。間引き手段106は、入力音声ベクトル105の要素値と合成音声ベクトル104の要素値の各々につき一定の間隔の間引き処理を行う。誤差評価手段109は、間引き入力音声ベクトル107及び間引き合成音声ベクトル108について二乗誤差の評価を行う。以上の処理が、励振ベクトル102毎に繰り返され、二乗誤差が最小(相関値が最大)となる合成音声ベクトルを生じさせたパラメータに基づき入力音声ベクトル105が符号化される。

本発明のブロック図



【特許請求の範囲】

【請求項1】 音声合成処理を行ってそれにより得られる合成音声ベクトルと入力音声ベクトルとの二乗誤差の評価処理を繰り返しながら該入力音声ベクトルの符号化を行う音声符号化方式において、前記入力音声ベクトルの要素値と前記合成音声ベクトルの要素値のそれぞれについて、一定の間隔の間引き処理を行う間引き手段と、該間引き手段によって前記入力音声ベクトル及び前記合成音声ベクトルの各要素値を間引いて得られる間引き入力音声ベクトル及び間引き合成音声ベクトルについて二乗誤差の評価処理を実行する誤差評価手段と、を有することを特徴とする音声符号化方式。

【請求項2】 コードブック記憶手段(101)から励振ベクトル(102)を選択し、該励振ベクトルに基づいて線形予測分析に基づく音声合成処理(103)を行い、それにより得られる合成音声ベクトル(104)と入力音声ベクトル(105)との二乗誤差の評価処理を繰り返しながら、該入力音声ベクトルの残差信号のベクトル量子化を行う処理を含むコード駆動線形予測方式に基づく音声符号化方式において、前記入力音声ベクトル(105)の要素値と前記合成音声ベクトル(104)の要素値のそれぞれについて、一定の間隔の間引き処理を行う間引き手段(106)と、該間引き手段によって前記入力音声ベクトル(105)及び前記合成音声ベクトル(104)の各要素値を間引いて得られる間引き入力音声ベクトル(107)及び間引き合成音声ベクトル(108)について二乗誤差の評価処理を実行する誤差評価手段(109)と、を有することを特徴とする音声符号化方式。

【請求項3】 コードブック記憶手段から励振ベクトルを選択し、該励振ベクトルに基づいて線形予測分析に基づく音声合成処理を行い、それにより得られる合成音声ベクトルと入力音声ベクトルとの二乗誤差の評価処理を聴覚特性に基づく重み付け処理を行った上で繰り返しながら、該入力音声ベクトルの残差信号のベクトル量子化を行う処理を含むコード駆動線形予測方式に基づく音声符号化方式において、前記入力音声ベクトルの要素値と前記合成音声ベクトルの要素値のそれぞれについて、一定の間隔の間引き処理を行う間引き手段と、該間引き手段によって前記入力音声ベクトル及び前記合成音声ベクトルの各要素値を間引いて得られる間引き入力音声ベクトル及び間引き合成音声ベクトルについて二乗誤差の評価処理を実行する誤差評価手段と、を有することを特徴とする音声符号化方式。

【請求項4】 前記誤差評価手段は、前記間引き入力音声ベクトル及び前記間引き合成音声ベクトルについて該両ベクトル間の相関値を演算し、該相関値が最大となる合成音声ベクトルを生じさせたパ

ラメータに基づいて前記入力音声ベクトルが符号化される、ことを特徴とする請求項1乃至3の何れか1項に記載の音声符号化方式。

【発明の詳細な説明】**【0001】**

【産業上の利用分野】本発明は、音声信号の情報圧縮伝送を行うための音声符号化方式に係り、更に詳しくは、コード駆動LPC符号化(CELP)方式などを用いた音声符号化方式における入力音声ベクトルと合成音声ベクトル間の誤差評価方式に関する。

【0002】

【従来の技術】音声は人間のコミュニケーションにおける重要な手段であることより、社会における音声情報の役割も非常に大きなものがある。OA化の進んだ現代においては、数字、文字と共に、電話での音声通信は必要不可欠な通信手段である。特に近年、デジタル信号処理技術とLSI技術に支えられたハードの驚異的な高性能化、小型化に伴い、デジタル通信路の拡大及びISDN(サービス総合デジタル網)などを用いた各種サービスのデジタル統合化が、我が国を初め諸外国で急速に進んでいる。

【0003】デジタル通信では、その基本尺度は使用回線の単位時間あたりの情報の伝送量(ビットレート)で決まるため、ネットワークの効率的かつ経済的利用と通信需要に対応するためにも、通信情報の圧縮技術の早急な開発が必要である。そして、人間の通信手段としての音声は、速報性・指示性・警報性といった機能を要求され、社会においては多くの場合、事務連絡・活動指示といった通信に使用されることを考えると、特に、デジタル移動無線システムや企業内通信システムなどにおいて、音声信号を4~16 Kbits/secの低・中ビットレートで高能率な情報圧縮を実現できる音声符号化方式が望まれている。

【0004】高能率音声符号化方式の代表的な方式として、音声信号を音声生成モデルに基いて分析してパラメータを抽出し、符号化を行う分析合成符号化方式がある。音声信号の大幅な情報圧縮を行うためには、音声信号からの音声情報の効率の良い抽出が必要であり、そのためには音声生成モデルの導入が不可欠である。音声生成モデルに基いた代表的かつ実用的な分析合成方式として線形予測モデルに基づく分析合成方式があり、LPC(線形予測符号化)方式と呼ばれている。

【0005】この方式は、音波形の標本値間に高い相関があることを利用し、現在の信号 x_j を過去のK個(10個程度)の標本の線形結合である予測値と、その時の誤差信号の和として表そうとするものであり、

【0006】**【数1】**

$$x_i = \sum_{j=1}^K a_j \cdot x_{i-j} + e_i$$

【0007】という式に帰着できる。ここで a_j は線形予測係数 (LPC係数)、 e_i は予測誤差信号と呼ばれ、この予測誤差信号 e_i の一定区間の平均2乗誤差を最小にするという条件でLPC係数を求めることを線形予測分析 (LPC分析) という。原音声信号 (符号化すべき音声信号) に対してLPC分析を行うことは、全極モデルに基づく音声生成システムを仮定したのと同価であり、その極は人間の主に声道の周波数特性におけるホルマントと呼ばれるスペクトルのピークに対応する。そして、LPC係数は、原音声信号の周波数特性が定常と考えられる一定区間 (例えば数十ミリ秒の区間) で、原音声信号を全極モデルによる音声生成モデルで近似した時のシステムパラメータとなっており、上記各区間毎にLPC分析によってLPC係数を演算しそれらを符号化することにより、音声の高効率符号化を実現できる。

【0008】ここで、上述の数1式において、右辺第1項は、K個のLPC係数 a_j をフィルタ係数とする音声合成フィルタの演算を表わしており、右辺第2項の予測誤差信号 e_i は、上記音声合成フィルタの出力 x_i を得るための系への入力、即ち音源と考えることができる。即ち、LPC方式において音声を合成するためには、音声合成フィルタを構成するためのLPC係数のほかに、音声合成フィルタへの入力である音源信号が必要となる。従って、LPC方式に基づいて音声を符号化する場合には、LPC係数を符号化すると共に、予測誤差信号 e_i を効率的に符号化する必要がある。

【0009】ここで、前述の数1式は、

【0010】

【数2】

$$e_i = x_i - \sum_{j=1}^K a_j \cdot x_{i-j}$$

【0011】と書き直すことができ、これより予測誤差信号 e_i は原音声信号 x_i を、LPC分析によってモデル化された線形システムの逆特性を有する逆フィルタに通すことによって得ることができる。この予測誤差信号は残差信号とも呼ばれる。

【0012】図7は、残差信号の時間域の波形の例を原音声信号と対比させて示した図である。LPC分析においてもし音声信号の予測が完全に行われれば、残差信号は値の小さい平坦なパワースペクトルを有する完全なランダム信号となる。しかし実際には、特に音声の大部分を占める有声音部分では、LPC分析による音声信号の予測が完全に行われず、残差信号には図7に示すように、周期的なパルスの時系列が現れる。また、パルス時系列の間には雑音時系列が現れる。一方、無声音部分では、残差信号は周波数的にほぼ白色な雑音時系列となる。

【0013】このように残差信号には、マクロ的にみれば有声音部分でパルス時系列の成分が含まれ、ミクロ的にみれば雑音時系列の成分が含まれる。そして、パルス時系列は人間の喉の声帯が振動することにより生ずるパルス列空気流に同期することが知られており、そのパルスの繰り返し周期はピッチ周期と呼ばれている。また、雑音時系列は、一定サンプル (例えば8 kHz サンプリングで64サンプル) 毎に、統計的に複数種類に分類された典型的な雑音時系列の候補の何れかによって近似的に表現可能である。

【0014】このような事実を利用して、移動無線等に適用可能な4~16 kbits/secの低・中ビットレートの音声伝送システムにおいて、残差信号を効率的かつ忠実に符号化することのできる音声符号化方式として、コード駆動LPC符号化 (CELP、以下同じ) 方式がある。

【0015】CELP方式では、送信装置において、次のような符号化処理が行われる。即ち、まず、伝送されるべき入力音声信号が所定サンプルに区切られたそれぞれの部分 (これをフレームと呼ぶ) について、LPC分析によって主に声道特性に対応するLPC係数が演算され、そのLPC係数の組が現在のフレームに対応する声道情報として符号化され、受信装置に伝送される。次に、入力音声信号からLPC係数で表現される成分を除いて得られる残差信号を表現する成分のうち、パルス時系列の成分と残りの雑音時系列の成分とが別々に符号化され、復号側の受信装置に伝送される。受信装置では、送信装置から伝送されてきた残差信号に関する各符号データに基づき、雑音時系列の成分が復号されると共にパルス時系列の成分が復号され、両者が加算されることによって残差信号が復号される。そして、送信装置から伝送されてきたLPC係数によってLPC合成フィルタが構成され、それに上述の復号された残差信号が入力されることによって、出力音声信号が合成される。

【0016】このように、CELP方式では、音声の声道特性に相当する成分をLPC係数として符号化すると共に、音声の音源特性に相当する残差信号をパルス時系列の成分と雑音時系列の成分とに分離して符号化することによって、高効率かつ高品質な音声の符号化を可能としている。

【0017】ここで、パルス時系列の候補を記憶したコードブックと、雑音時系列の候補を記憶したコードブックのそれぞれが、送信装置と受信装置の両方に用意される。送信装置と受信装置とは、それぞれ同じパルス時系列のコードブック、同じ雑音時系列のコードブックが用意される。

【0018】そして、送信装置では、フレーム単位の残差信号を最適に近似するパルス時系列がパルス時系列のコードブック中の候補の中から探索され、コードブック上での当該パルス時系列の位置を示す番号 (インデックス) が受信装置に伝送される。この時、最適なパルス時系列の振幅を残差信号の振幅に適合させるための最適な

ゲインも演算・符号化されて受信装置に伝送される。次に、残差信号から上述の最適なパルス時系列に最適なゲインを乗算した成分が除かれ、フレーム単位の残差誤差信号が得られる。そして、このフレーム単位の残差誤差信号を最適に近似する雑音時系列が、パルス時系列の候補の探索の場合と同様にして、雑音時系列のコードブック中の候補の中から探索され、そのインデックスが受信装置に伝送される。また、最適な雑音時系列の振幅を上記残差誤差信号の振幅に適合させるための最適なゲインも演算・符号化されて受信装置に伝送される。

【0019】上述の送信装置側の動作に対応して、受信装置では、送信装置から送られてきた雑音時系列についてのインデックスで雑音時系列のコードブックが参照され、対応する雑音時系列が読み出される。そして、その雑音時系列に送信装置から送られてきた最適なゲインが乗算されることにより、残差誤差信号が復号される。続いて、送信装置から送られてきたパルス時系列についてのインデックスでパルス時系列のコードブックが参照され、対応するパルス時系列が読み出される。そして、そのパルス時系列に送信装置から送られてきた最適なゲインが乗算され、この結果得られたパルス時系列信号に上述の残差誤差信号が加算されることにより、残差信号が復号される。

【0020】ここで、送信装置において、パルス時系列のコードブックから残差信号を近似する最適なパルス時系列と最適なゲインを探索し、また、雑音時系列のコードブックから残差誤差信号を近似する最適な雑音時系列と最適なゲインを探索するCELP処理につき、図8の原理構成図を用いて説明する。図8において、801が上述の前者の処理を行う部分、802が後者の処理を行う部分である。

【0021】まず、パルス時系列のコードブック中の候補の中からフレーム単位の残差信号を最適に近似するパルス時系列を探索するためには、単純には入力音声信号に対するLPC分析に基づいて得られる各フレームの残差信号(数2式参照)と、コードブック中の各パルス時系列の候補との誤差を計算し、それが最小となるものを選択すればよい。しかし、実際には人間の聴覚上の誤差が最小となるパルス時系列を選択すべきであるため、音声信号のレベルで誤差評価が行われる。しかもその場合に、そのままの音声信号ではなく、人間の聴覚に合うように周波数特性に重み付けがなされた音声信号のレベルで誤差評価が行われる。

【0022】即ち、図8の部分801において、まず、聴覚重み付け部803は、1フレーム分の入力音声ベクトル s (複数サンプルからなる入力音声信号)に対して、人間の聴覚に合うようにその周波数特性に重み付けがなされる処理を実行することによって、重み付入力音声ベクトル x (複数サンプルからなる重み付入力音声信号)を出力する。

【0023】一方、パルスベクトルコードブック807(パルス時系列のコードブック)には、それぞれ複数サンプルからなる複数組のパルスベクトル(パルス時系列)が記憶されている。ここで、1組のパルスベクトルを $p(m)$ とする。 m はパルスベクトルコードブック807上にある複数組のパルスベクトルのインデックスを表わし、組の数を α とすれば $1 \leq m \leq \alpha$ である。

【0024】今、誤差評価部804が読出し部806に対して $m=1$ である1組目のパルスベクトルを読み出すよう指示する。これにより、読出し部806によりパルスベクトルコードブック807(パルス時系列のコードブック)から読み出された1組目のパルスベクトル(パルス時系列) $p(1)$ が、重み付短期予測部805に入力される。

【0025】重み付短期予測部805は、LPC係数 a_j ($1 \leq j \leq K$)で構成されるLPC合成フィルタの処理を聴覚重み付けを考慮した上で実行することにより、1組目のパルスベクトル $p(1)$ に対応する合成音声ベクトル $y(1)$ (複数サンプルからなる合成音声信号)を演算する。このLPC係数 a_j ($1 \leq j \leq K$)は、前述した入力音声ベクトル s に対して特には図示しない K 次のLPC分析処理が実行されることによって得られる。

【0026】誤差評価部804は、上述の1組目のパルスベクトル $p(1)$ に対応する合成音声ベクトル $y(1)$ と重み付入力音声ベクトル x との誤差を計算する。そして、誤差評価部804は、1組目の誤差評価を終了したら、 m の値を更新して2とし、読出し部806に、パルスベクトルコードブック807から $m=2$ に対応する次の組のパルスベクトル $p(2)$ を読み出させる。

【0027】このようにして、誤差評価部804は、 m の値を1から α まで更新しながら、パルスベクトルコードブック807から順次読み出されるパルスベクトル $p(m)$ に対応する合成音声ベクトル $y(m)$ について、聴覚重み付け部803で求められている同じ重み付入力音声ベクトル x との誤差を演算し、全てのパルスベクトル $p(m)$ ($1 \leq m \leq \alpha$)についての誤差評価を終了した時点で、誤差が最も小さかったパルスベクトルに対応するインデックス m の値 m_{opt} と最適なゲイン g を出力する。

【0028】以上のように、音声信号(音声ベクトル)を合成しながら誤差評価を行う方式は、一般に、合成による分析手法(Analysis-by-Synthesis:A-b-S法)と呼ばれている。A-b-S法では、入力音声信号から残差信号を直接求める必要はないが、コードブックの探索を行う毎に音声合成処理を実行する必要がある。

【0029】ここで、誤差評価部804におけるゲインを考慮した誤差評価方式について説明する。まず、重み付入力音声ベクトル x と合成音声ベクトル $y(m)$ との誤差電力 E_{dl} は、次式で表わされる。

【0030】

【数3】

$$E_d = \sum_{i=1}^N (x_i - g \cdot y_i(m))^2$$

【0031】ここで、合成音声ベクトル $y(m)$ は重み付短期予測部805において、

【0032】

【数4】

$$y_i(m) = p_i(m) + \sum_{j=1}^K a_j \cdot y_{i-j}(m)$$

【0033】なるフィルタ演算で合成される。なお、重

$$dE_d / dg = 0$$

$$2 \sum_{i=1}^N (x_i - g \cdot y_i(m)) \cdot y_i(m) = 0$$

$$\therefore g = \frac{\sum_{i=1}^N x_i \cdot y_i(m)}{\sum_{i=1}^N |y_i(m)|^2}$$

【0036】この時の誤差電力 E_d は、数5式を数3式に代入することにより、次式のようになる。

$$E_d = \sum_{i=1}^N |x_i|^2 -$$

み付入力音声ベクトル x と合成音声ベクトル $y(m)$ の各添え字 i は、 N サンプルからなる1フレーム(1組)分の音声サンプルの各要素の番号を示しており、 $1 \leq i \leq N$ である。

【0034】上述の数3式の誤差電力 E_d を最小にするゲイン g は、数3式を g で微分して零とおくことにより得られる。即ち、次式の通りである。

【0035】

【数5】

【0037】

【数6】

$$\frac{\left\{ \sum_{i=1}^N x_i \cdot y_i(m) \right\}^2}{\sum_{i=1}^N |y_i(m)|^2}$$

【0038】数6式の右辺第1項は、重み付入力音声ベクトル x の電力を表わしており、各パルスベクトル $p(m)$ の組が変更されることによって合成音声ベクトル $y(m)$ の組が変更されても一定である。従って、誤差電力 E_d を最小にするパルスベクトル $p(m)$ の組は、数6式の右辺第2項、即ち、

【0039】

【数7】

$$A = \frac{\left\{ \sum_{i=1}^N x_i \cdot y_i(m) \right\}^2}{\sum_{i=1}^N |y_i(m)|^2}$$

【0040】を最大にするパルスベクトル $p(m)$ の組として求めることができる。この数7式は、重み付入力音声ベクトル x と合成音声ベクトル $y(m)$ の一種の相関を求める演算である。従って、数7式を最大にするパルスベクトル $p(m)$ を求めることは、2つの音声ベクトルの相関が最大なものを求めることにほぼ等価である。

【0041】以上の誤差評価の原理に基づいて、図8の誤差評価部804は、各パルスベクトル $p(m)$ の組に対

応する合成音声ベクトル $y(m)$ について数7式を演算し、その値 A が最終的に最も大きくなる組に対応するインデックス m の値を m_{opt} として出力する。即ち、数7式の値を最大にする最適なパルスベクトル及び合成音声ベクトルは、 $p(m_{opt})$ 及び $y(m_{opt})$ である。

【0042】そして、誤差評価部804は、最適合成音声ベクトル $y(m_{opt})$ を用いて数5式を演算することにより、最適ゲイン g を出力する。以上の処理の結果、誤差評価部804から乗算部808に上述の最適合成音声ベクトル $y(m_{opt})$ が出力される。そして、乗算部808において $y(m_{opt})$ に最適ゲイン g が乗算され、更に、減算部809において上述の乗算結果が重み付入力音声ベクトル x_i から減算されることにより、重み付入力音声誤差ベクトル e_x が得られる。

【0043】次に、雑音時系列のコードブック中の候補の中からフレーム単位の残差誤差信号を最適に近似する雑音時系列を探索する場合にも、聴覚重み付けがなされた音声信号のレベルで誤差評価が行われる。

【0044】即ち、図8の部分802において、雑音ベクトルコードブック813(雑音時系列のコードブック)には、それぞれ複数サンプルからなる複数組の雑音ベクトル(雑音時系列)が記憶されている。ここで、1

組の雑音ベクトルを $c(n)$ とする。 n は雑音ベクトルコードブック 813 上にある複数組の雑音ベクトルのインデックスを表わし、組の数を β とすれば $1 \leq n \leq \beta$ である。

【0045】誤差評価部 810 による制御動作は、前述した誤差評価部 804 の場合とほとんど同じである。即ち、誤差評価部 810 は、 n の値を 1 から β まで更新しながら、雑音ベクトルコードブック 813 から順次読み出される雑音ベクトル $c(n)$ に対応して重み付短期予測部 811 で合成される合成音声誤差ベクトル $e_y(n)$ について、重み付入力音声誤差ベクトル e_x との誤差を演算し、全ての雑音ベクトル $c(n)$ ($1 \leq n \leq \beta$) につい

$$e_{y_i}(n) = c_i(n) + \sum_{j=1}^K a_j \cdot e_{y_{i-j}}(n)$$

【0049】

$$dE_d / dg = 0$$

$$2 \sum_{i=1}^N (e_{x_i} - g \cdot e_{y_i}(n)) \cdot e_{y_i}(n) = 0$$

$$\therefore g = \frac{\sum_{i=1}^N e_{x_i} \cdot e_{y_i}(n)}{\sum_{i=1}^N |e_{y_i}(n)|^2}$$

【0050】

$$E_d = \frac{\sum_{i=1}^N |e_{x_i}|^2}{\sum_{i=1}^N |e_{y_i}(n)|^2} = \frac{\left\{ \sum_{i=1}^N e_{x_i} \cdot e_{y_i}(n) \right\}^2}{\sum_{i=1}^N |e_{y_i}(n)|^2}$$

【0051】

【数12】

$$A = \frac{\left\{ \sum_{i=1}^N e_{x_i} \cdot e_{y_i}(n) \right\}^2}{\sum_{i=1}^N |e_{y_i}(n)|^2}$$

【0052】即ち、誤差評価部 810 は、読出し部 812 によって雑音ベクトルコードブック 813 から読み出される各雑音ベクトル $c(n)$ の組に対応して重み付短期予測部 811 で数 9 式に基づいて合成される合成音声誤差ベクトル $e_y(n)$ について数 12 式を演算し、その値 A が最終的に最も大きくなる組に対応するインデックス n の値を n_{opt} として出力する。更に、誤差評価部 810 は、最適合成音声誤差ベクトル $e_y(n_{opt})$ を用いて数 10 式を演算することにより、最適ゲイン b を出力する。

ての誤差評価を終了した時点で、誤差が最も小さかった雑音ベクトルに対応するインデックス n の値 n_{opt} と最適なゲイン b を出力する。

【0046】そして、誤差評価部 810 におけるゲインを含めた誤差評価は、前述した数 3 式～数 7 式に対応する以下の数 8 式～数 12 に基づいて行われる。

【0047】

【数8】

$$E_d = \sum_{i=1}^N (e_{x_i} - b \cdot e_{y_i}(n))^2$$

【0048】

【数9】

【数10】

【数11】

$$\left\{ \sum_{i=1}^N e_{x_i} \cdot e_{y_i}(n) \right\}^2$$

$$\sum_{i=1}^N |e_{y_i}(n)|^2$$

【0053】以上に示される送信装置側の動作に対応して、受信装置では、送信装置から送られてきたインデックス n_{opt} で図 8 の 813 と全く同じ構成の受信装置内の雑音ベクトルコードブックが参照され、対応する最適雑音ベクトル $c(n_{opt})$ が読み出される。そして、それに送信装置から送られてきた最適ゲイン b が乗算されることにより、残差誤差信号が復号される。続いて、送信装置から送られてきたインデックス m_{opt} で図 8 の 807 と全く同じ構成の受信装置内のパルスベクトルコードブックが参照され、対応する最適パルスベクトル $p(m_{opt})$ が読み出される。そして、それに送信装置から送られてきた最適ゲイン g が乗算され、この結果得られたパルスベクトルに上述の残差誤差信号が加算されることにより、残差信号が復号される。

【0054】

【発明が解決しようとする課題】以上に示した A-b-S 法を基本とする CELP 方式の従来例においては、コードブックの探索を行う毎に音声合成処理を実行する必要が

あるため、最適なインデックス m_{opt} 、 n_{opt} 及び最適ゲイン g 、 b を得るまでに、非常に多くの演算量を必要とする。

【0055】特に、優れた合成音声の品質を得るためには、図8の雑音ベクトルコードブック813のサイズ β を大きくする必要がある。これは、次のような理由による。即ち、残差信号を近似すべきパルスベクトルの種類はそれほど多くはないのに対して、残差信号からパルスベクトルの成分を除いた残差誤差信号は、周波数特性が白色に近く、時間領域のベクトルパターンとしても非常に多くの種類のパターンが出現し得る。従って、このような残差誤差信号を正確に近似するためには、多くのパターンの雑音ベクトルが必要となるためである。

【0056】このように、大きなサイズの雑音ベクトルコードブック813が必要となる結果、図8の誤差評価部810が数12式を演算する回数が増加し、全体的に膨大な演算量となってしまいう問題点を有している。

【0057】上述のような問題点は、CELP方式に限られることなく、例えば残差信号の符号化方式の1つであるマルチパルス符号化方式において、A-b-S法に基づいて入力音声と合成された音声の二乗誤差の評価を繰り返す処理などにおいても、同様に発生する。

【0058】本発明は、合成音声の品質の劣化を最小限に抑えつつ、誤差評価時の演算量を削減することによって、音声符号化処理の高速化を図ることを目的とする。

【0059】

【課題を解決するための手段】図1は、本発明のブロック図である。本発明は、音声合成処理を行ってそれにより得られる合成音声ベクトルと入力音声ベクトルとの二乗誤差の評価処理を繰り返しながら入力音声ベクトルの符号化を行う音声符号化方式を前提とする。より具体的には、例えばコード駆動線形予測方式に基づく音声符号化方式が適用される。この方式では、まず、入力音声ベクトル105が線形予測分析されてその線形予測係数が符号化される。この処理と共に、コードブック記憶手段101から励振ベクトル102が選択され、その励振ベクトルと上述の線形予測係数に基づいて音声合成処理手段103において音声合成処理が実行され、それにより得られる合成音声ベクトル104と入力音声ベクトル105との二乗誤差の評価処理が繰り返されながら、入力音声ベクトル105の残差信号のベクトル量子化が行われる。その他、音声合成方式として、残差信号の符号化方式の1つであるマルチパルス符号化方式なども採用できる。

【0060】そして、まず、入力音声ベクトル105の要素値と合成音声ベクトル104の要素値のそれぞれについて、一定の間隔の間引き処理を行う間引き手段106を有する。

【0061】次に、間引き手段106によって入力音声

ベクトル105及び合成音声ベクトル104の各要素値を間引いて得られる間引き入力音声ベクトル107及び間引き合成音声ベクトル108について二乗誤差の評価処理を実行する誤差評価手段109を有する。同手段は、例えば、間引き入力音声ベクトル107及び間引き合成音声ベクトル108について両ベクトル間の相関値を演算する。

【0062】上述の二乗誤差の評価処理が繰り返されることにより、二乗誤差が最小（相関値が最大）となる合成音声ベクトルを生じさせたパラメータに基づいて入力音声ベクトル105が符号化される。

【0063】上述の本発明の構成において、二乗誤差の評価処理は聴覚特性に基づく重み付け処理を行った上で実行されるように構成することができる。具体的には、入力音声ベクトル105について聴覚重み付けが行われ、また、音声合成処理手段103が聴覚重み付けを考慮した音声合成処理を行うことによって聴覚重み付けされた合成音声ベクトル104が生成される。そして、これら両ベクトルについて間引き手段106による間引き処理、誤差評価手段109による誤差評価が実行される。

【0064】

【作用】音声合成処理を行ってそれにより得られる合成音声ベクトル104と入力音声ベクトル105との二乗誤差の評価処理を繰り返しながら入力音声ベクトル105の符号化を行う音声符号化方式では、1回の二乗誤差の評価処理において、合成音声ベクトル104と入力音声ベクトル105の各要素間の乗算演算等が必要となる。そして、例えばコード駆動線形予測方式において、復号音声の音質を向上させるためにコードブック記憶手段101に記憶される励振ベクトル102の種類が増えると、その全ての励振ベクトルについて二乗誤差の評価処理を行わねばならないため、全体的に膨大な演算量が必要となる。

【0065】そこで、本発明では、音声信号は隣接する数サンプル間では近接相関が高いことを利用して、間引き手段106によって入力音声ベクトル105及び合成音声ベクトル104の各要素値を間引いて得られる間引き入力音声ベクトル107及び間引き合成音声ベクトル108について誤差評価手段109で誤差評価が行われる。

【0066】これにより、復号装置側での合成音声の品質の劣化を最小限に抑えつつ、誤差評価のための演算量を数分の1にすることができ、音声符号化装置を高速化することができ、回路規模の縮小が可能となる。

【0067】

【実施例】以下、図面を参照しながら本発明の実施例につき詳細に説明する。本発明は、大きく分けて、図2のLPC符号化部と図3の残差符号化部とで構成される。

【0068】なお、本発明の実施例では、後述する図3

の誤差評価部310の構成と動作に特徴がある。

<LPC符号化部の説明>図2は、LPC符号化部の構成図である。

【0069】サンプリング周波数8 kHz でサンプリングされた入力音声信号は、バッファ部201に、入力音声ベクトル s として例えば1フレーム=320サンプルずつ保持される。そして、320サンプルの入力音声ベクトル s に対する符号化処理が、320サンプリング周期の間にリアルタイムで実行される。

【0070】LPC分析部202は、上述の320サンプルの入力音声ベクトル s に対してLPC分析を実行し、1組のLPC係数 a_j ($1 \leq j \leq K$) を出力する。LPC分析の具体的手法としては、例えばPARCOR分析法、LSP分析法等が適用でき、LPC分析の次数 K は例えば10次である。

【0071】入力音声ベクトル s の320サンプル毎にLPC分析部202で求まる1組の10次のLPC係数 a_j は、符号化部203で符号化された後、特に図示しない音声復号装置に伝送される。

【0072】また、この符号化されたLPC係数 a_j は、復号部204において音声復号装置における復号処理と同じ処理によって復号され、後述する残差符号化部での処理に使用される。このように、復号したLPC係数 a_j が用いられるのは、LPC係数 a_j を符号化するときには発生する量子化誤差を残差信号に含めてしまいそのような残差信号を符号化することによって、LPC係数 a_j の符号化により発生する復号時の音声の品質劣化を最小限に抑えるためである。

<残差符号化部の説明>次に、図3は、残差符号化部の構成図である。

【0073】図3において、301~313の参照番号が付された各部は、その番号順に、「従来の技術」の項で前述した図8のCELP方式の原理構成における801~813の各部に対応している。そして、図3においては、その基本的な構成は図8のCELP方式の原理構成と同様だが、CELP方式を実際の残差符号化装置として実現するための構成が付加されている。

【0074】即ち、まず、図2のLPC符号化部のバッファ部201から出力されている1フレーム(=320サンプル)分の入力音声ベクトル s に対して、聴覚重み付け部303で、人間の聴覚に合うようにその周波数特性に重み付けがなされ、そこから得られる64サンプルの重み付入力音声ベクトル x は、バッファ部314に保持される。そして、バッファ部314に保持された64サンプルの重み付入力音声ベクトル x に対して、301及び302の部分の処理が連続的に実行される。

【0075】ここで、聴覚重み付け部303、301及び302の各部分における1サブフレーム=64サンプルに対する処理が5サブフレーム分連続して実行されることにより、1フレーム=320サンプルの入力音声ベ

クトル s に対する処理が実行されたことになる。

【0076】従って、符号化装置から特に図示しない復号装置に対しては、図2のLPC符号化部から320サンプルに1組の割合で10次のLPC係数 a_j が出力され、また、図3の残差符号化部から320サンプルに5組の割合で最適インデックス m_{opt} と最適ゲイン g の組が、同様に、320サンプルに5組の割合で最適インデックス n_{opt} と最適ゲイン b の組が出力されることになる。

第1段目の残差符号化処理の説明

より具体的には、まず、301の部分において、バッファ部314に保持された64サンプルの重み付入力音声ベクトル x の各サブフレームについて、以下のような誤差評価処理が実行される。

【0077】図3の誤差評価部304の構成を図4に示す。誤差評価部304は、評価値演算部410、最大評価値記憶部402、合成音声ベクトル記憶部403、インデックス記憶部404、インデックスカウンタ部405及びゲイン演算部406から構成される。以下、図3と図4を用いて説明する。

【0078】始めに、アダプティブコードブック307は、図8のパルスベクトルコードブック807と同様であり、そこには α 組のパルスベクトル $p(m)$ ($1 \leq m \leq \alpha$) が記憶されている。ここで、各組のパルスベクトルの要素の数(サンプル数)は上述の例では64サンプル、即ち、 $p(m) = p_i(m)$ ($1 \leq i \leq 64$) である。また、本実施例においては、パルスベクトル $p(m)$ は、後述する更新部327によって更新される。

【0079】今、誤差評価部304内の評価値演算部401が、インデックスカウンタ部405に初期値1をセットし、インデックスカウンタ部405から読出し部306に対して $m=1$ が出力される。これにより、読出し部306は、アダプティブコードブック307から1組目のパルスベクトル $p(1)$ を読み出し、それを重み付短期予測部305に出力する。

【0080】重み付短期予測部305は、LPC係数 a_j ($1 \leq j \leq K$) で構成される前述した数4式に基づくLPC合成フィルタの処理を、聴覚重み付けを考慮した上で実行することにより、1組目のパルスベクトル $p(1)$ に対応する合成音声ベクトル $y(1)$ を演算する。このLPC係数 a_j ($1 \leq j \leq K$) は、図2の符号化部203から供給されている。

【0081】誤差評価部304内の評価値演算部401は、上述の1組目のパルスベクトル $p(1)$ に対応する合成音声ベクトル $y(1)$ ($=y_i(1)$, $1 \leq i \leq 64$) とバッファ部314から読み込んだ1サブフレーム(=64サンプル)分の重み付入力音声ベクトル x ($=x_i$, $1 \leq i \leq 64$) について、前述した数7式の評価値 A を演算する。

【0082】評価値演算部401は、1組目の誤差評価

を終了したら、評価値A、1組目の合成音声ベクトル y (m) 及びインデックス $m=1$ を、それぞれ最大評価値記憶部402、合成音声ベクトル記憶部403及びインデックス記憶部404に記憶する。

【0083】続いて、インデックスカウンタ部405は、評価値演算部401からの指示に基づいて、カウンタ値であるインデックス m の値を1から2にインクリメントする。これにより、読出し部306は、アダプティブコードブック307から2組目のパルスベクトル p (2)を読み出し、それを重み付短期予測部305に出力する。

【0084】評価値演算部401は、上述のようにインデックスカウンタ部405に対して m の値を1から α まで更新させながら、次のような制御処理を行う。即ち、評価値演算部401は、アダプティブコードブック307から読み出されるパルスベクトル p (m) に対応する合成音声ベクトル y (m) ($=y_i(m), 1 \leq i \leq 64$) とバッファ部314から出力されている1サブフレーム分の重み付入力音声ベクトル x ($=x_i, 1 \leq i \leq 64$) とにつき、前述した数7式の評価値Aを演算する。そして、評価値演算部401は、その評価値Aを最大評価値記憶部402に記憶されている評価値と比較し、今回演算された評価値Aの方が大きければ、最大評価値記憶部402の内容を今回の評価値Aで置き換え、また、合成音声ベクトル記憶部402及びインデックス記憶部404の各内容を、それぞれ今回の合成音声ベクトル y (m) 及び今回のインデックス m の値で置き換える。今回演算された評価値Aが最大評価値記憶部402に記憶されている評価値以下の場合には、上述の各置き換えは行われない。

【0085】このようにして、全てのインデックス値 m ($1 \leq m \leq \alpha$) に対する誤差評価が終了した時点において、インデックス記憶部404には数7式の評価値Aを最大にしたパルスベクトルのインデックス値 $m=m_{opt}$ が記憶されており、また、合成音声ベクトル記憶部403にはそのときの最適合成音声ベクトル y (m_{opt}) が記憶されていることになる。

【0086】この後、評価値演算部401は、バッファ部314から出力されている重み付入力音声ベクトル x ($=x_i, 1 \leq i \leq 64$) と、合成音声ベクトル記憶部403から読み出した最適合成音声ベクトル y (m_{opt}) ($=y_i(m_{opt}), 1 \leq i \leq 64$) をゲイン演算部406に出力する。ゲイン演算部406は、上述の2組のベクトルに基づいて前述した数5式を演算することによって最適ゲイン g を出力する。そして、最適ゲイン g は、符号化部318で符号化される。

【0087】以上のようにして、評価値演算部401からは最適インデックス m_{opt} が、また、符号化部318からは符号化された最適ゲイン g が、それぞれ特には図示しない音声復号装置へ送出される。

【0088】なお、評価値演算部401は、インデックス記憶部404から読み出した最適インデックス m_{opt} を読出し部306へ出力する。読出し部306は、最適インデックス m_{opt} に対応する最適パルスベクトル p (m_{opt}) ($=p_i(m_{opt}), 1 \leq i \leq 64$) を、アダプティブコードブック307から読み出し、バッファ部322へ出力する。このバッファ部322のデータは、アダプティブコードブック307の後述する更新処理に使用される。

【0089】上述の処理に続き、評価値演算部401は、合成音声ベクトル記憶部403から読み出した最適合成音声ベクトル y (m_{opt}) ($=y_i(m_{opt}), 1 \leq i \leq 64$) を図3の乗算部308へ出力する。また、符号化部318によって符号化された最適ゲイン g は、復号部319において特には図示しない音声復号装置における復号処理と同じ復号処理によって復号された後に、乗算部308に入力される。このように、復号した最適ゲインが用いられるのは、復号部204で説明したのと同様、符号化による量子化誤差の影響を除去するためである。そして、乗算部308において、最適合成音声ベクトル y (m_{opt}) に最適ゲイン g を復号部319で復号した値が乗算され、更に、減算部309において上述の乗算結果がバッファ部314から出力されている1サブフレーム分の重み付入力音声ベクトル x ($=x_i, 1 \leq i \leq 64$) から減算されることにより、1サブフレーム分の重み付入力音声誤差ベクトル e_x ($=e_{x_i}, 1 \leq i \leq 64$) が得られる。

第2段目の残差符号化処理の説明

次に、図3の302の部分において、バッファ部316に保持された64サンプルの重み付入力音声誤差ベクトル e_x の各サブフレームについて、以下のような誤差評価処理が実行される。

【0090】図3の誤差評価部310の構成を図5に示す。誤差評価部310は、図4の誤差評価部304における401～406の構成に対応して、評価値演算部501、最大評価値記憶部502、合成音声ベクトル記憶部503、インデックス記憶部504、インデックスカウンタ部505及びゲイン演算部506を有するが、更に、間引き部507及び508を有することが本実施例の最も大きな特徴である。以下、図3と図5を用いて説明する。

【0091】始めに、ストカスティックコードブック313は、図8の雑音ベクトルコードブック813と同様であり、そこには β 組の雑音ベクトル c (n) ($1 \leq n \leq \beta$) が記憶されている。ここで、各組の雑音ベクトルの要素の数(サンプル数)は本実施例では64サンプル、即ち、 $c(n) = c_i(n)$ ($1 \leq i \leq 64$) である。

【0092】そして、評価値演算部501は、図4の401と同様に、インデックスカウンタ部505に対して n の値を1から β まで更新させながら、以下のような制

御処理を行う。即ち、評価値演算部501は、読出し部512によってストカスティックコードブック313から読み出される各雑音ベクトル $c(n)$ の組に対応して重み付短期予測部311で数9式に基づいて合成されバッファ部317を介して入力される合成音声誤差ベクトル $e_y(n)$ ($=e_{y_i}(n)$, $1 \leq i \leq 64$)と、バッファ部316から出力されている1サブフレーム分の重み付入力音声誤差ベクトル e_x ($=e_{x_i}$, $1 \leq i \leq 64$)とにつき、誤差の評価値を演算する。そして、評価値演算部501は、その評価値を最大評価値記憶部502に記憶されている評価値と比較して、今回演算された評価値の方が大きければ、最大評価値記憶部502の内容を今回の評価値で置き換え、また、合成音声誤差ベクトル記憶部503及びインデックス記憶部504の各内容を、それぞれ今回の合成音声誤差ベクトル $e_y(n)$ 及び今回のインデックス n の値で置き換える。今回演算された評価値が最大評価値記憶部502に記憶されている評価値以下の場合には、上述の各置き換えは行われない。

【0093】ここで、評価値演算部501における評価値の演算式は、図4の評価値演算部401における評価値の演算式である数7式が数3式～数6式に基づいて決定されたのに対応して、前述した数8式～数11式に基づいて決定される。そして、これらの式に基づいて基本的に導出される評価値の演算式は、従来例でも使用していた数12式である。本実施例では、この数12式を以下に示す数13式のように変形した点が大きな特徴である。

【0094】

【数13】

$$A = \frac{\left\{ \sum_{i=1}^{N(\text{step } M)} e_{x_i} \cdot e_{y_i}(n) \right\}^2}{\sum_{i=1}^{N(\text{step } M)} e_{y_i}(n)^2}$$

【0095】即ち従来例では、数12式のように、1サブフレーム分の重み付入力音声誤差ベクトル e_x ($=e_{x_i}$, $1 \leq i \leq 64$)と1組分の合成音声誤差ベクトル $e_y(n)$ ($=e_{y_i}(n)$, $1 \leq i \leq 64$)の、それぞれの全ての要素値（信号サンプル値）を使用して評価値Aが演算された。これに対して本実施例では、数13式のように、各ベクトルの要素値が一定間隔Mで間引かれながら評価値Aが演算される。一般に、音声信号及び合成音声信号は、8乃至10次程度の近接相関を有していることが良く知られている。そこで、数13式のように間引き処理を行いながら評価値Aを演算しても、数12式のように間引き処理を行わないで評価値Aを演算した場合に比較して、評価値Aは大きくは変化しないことが期待されるのである。ここで、数13式の“N(step M)”は、ベクトルの要素値を示す番号iが1から値Mずつ値

Nまでインクリメントされることを示している。

【0096】本実施例では、M=4程度までの数13式に基づく間引き演算では、数12式の演算を行った場合に比較して、復号された音声において、ほとんど遜色のないS/N（若しくはセグメンタルS/N）を得られることが、実験的に確認されている。

【0097】ここで、数12式及び数13式のそれぞれによって評価値Aを演算した場合の演算量を比較すると、図6のようになる。これより、例えば1サブフレームのサンプル数N=64、間引き間隔M=4とした場合に、数13式で評価値Aを演算した方が、約1/4の加算及び乗算の回数で済むことがわかる。

【0098】以上の原理に基づいて、図3又は図4の誤差評価部310では、次のような評価値Aの演算処理が行われる。即ち、まず、バッファ部316から出力される1サブフレーム分の重み付入力音声誤差ベクトル e_x ($=e_{x_i}$, $1 \leq i \leq 64$)は、間引き部507においてサンプル間隔Mで間引かれて、評価値演算部501に入力される。また、バッファ部317から出力される1組の合成音声誤差ベクトル $e_y(n)$ ($=e_{y_i}(n)$, $1 \leq i \leq 64$)は、間引き部508においてやはりサンプル間隔Mで間引かれて、評価値演算部501に入力される。これらの処理は、数13式において、ベクトルの要素値を示す番号iが1から値Mずつ値Nまでインクリメントされる処理に相当する。そして、評価値演算部501は、数13式における上述の動作以外の残りの演算処理を実行するのである。

【0099】なお、合成音声誤差ベクトル記憶部503には、間引かれた合成音声誤差ベクトルではなく、間引かれていない合成音声誤差ベクトル $e_y(n)$ ($=e_{y_i}(n)$, $1 \leq i \leq 64$)が記憶される。これは、後述する最適ゲインの演算は間引き処理なしに実行された方が、音声復号装置で合成される音声の音質を向上させることができ、また、最適ゲインの演算は1サブフレームに1回実行されるだけなので間引き処理をしなくても演算量はそれほど増大しないからである。

【0100】以上のような評価値Aの演算処理を基本として、全てのインデックス値 n ($1 \leq n \leq \beta$)に対する誤差評価が終了した時点で、インデックス記憶部504には数13式の評価値Aを最大にした雑音ベクトルのインデックス値 $n=n_{\text{opt}}$ が記憶されており、また、合成音声誤差ベクトル記憶部503にはそのときの最適合成音声誤差ベクトル $e_y(n_{\text{opt}})$ ($=e_{y_i}(n)$, $1 \leq i \leq 64$)が記憶されていることになる。

【0101】この後、評価値演算部501は、バッファ部316から出力されている重み付入力音声誤差ベクトル e_{x_i} ($=e_{x_i}$, $1 \leq i \leq 64$)と、合成音声ベクトル記憶部403から読み出した最適合成音声誤差ベクトル $e_y(n_{\text{opt}})$ ($=e_{y_i}(n_{\text{opt}})$, $1 \leq i \leq 64$)をゲイン演算部506に出力する。ゲイン演算部506

は、上述の2組のベクトルに基づいて従来例と同様の数10式を演算することによって最適ゲイン b を出力する。そして、最適ゲイン b は、符号化部320で符号化される。

【0102】以上のようにして、評価値演算部501からは最適インデックス n_{opt} が、また、符号化部320からは符号化された最適ゲイン b が、それぞれ特には図示しない音声復号装置へ送出される。

アダプティブコードブック307の更新処理の説明

最後に、アダプティブコードブック307の更新処理について説明する。

【0103】評価値演算部501は、1サブフレーム(=64サンプル)の上述した誤差評価処理を終了した後に、インデックス記憶部504から読み出した最適インデックス n_{opt} を読み出し部312へ出力する。読み出し部312は、この最適インデックス n_{opt} に対応する最適雑音ベクトル $c(n_{opt})$ (= $c_i(n_{opt})$, $1 \leq i \leq 64$)をストカスティックコードブック313から読み出して、乗算部324へ出力する。一方、バッファ部322から乗算部323へは、64サンプルの最適パルスベクトル $p(m_{opt})$ (= $p_i(m_{opt})$, $1 \leq i \leq 64$)のうちの上記64サンプルの最適雑音ベクトル $c(n_{opt})$ に対応する部分が出力される。乗算部323では最適パルスベクトル $p(m_{opt})$ に最適ゲイン g を復号部319で復号した値が乗算され、乗算部324では最適雑音ベクトル $c(n_{opt})$ に最適ゲイン b を復号部321で復号した値が乗算される。ここで、復号した最適ゲインが用いられるのは、復号部204で説明したのと同様、符号化による量子化誤差の影響を除去するためである。そして、これら2つの乗算結果が加算部325で加算され、バッファ部326に格納される。

【0104】このようにしてバッファ部326に得られる64サンプル分の更新データは、現在の64サンプルの区間において特には図示しない音声復号装置において得られるであろう最適な残差信号を表現していることになる。そして、この最適な残差信号が、更新部327によって、アダプティブコードブック307に新たなパルスベクトルの組として書き込まれる。このとき、コードブック上の最も古い1組のパルスベクトルが捨てられる。

【0105】以上のようにアダプティブコードブック307には、図3の301での処理における現フレームまでに求まった最適な残差信号の組が新しい順に記憶され、これらのデータが次のサブフレームにおけるパルスベクトルの探索に使用される。これは、残差信号は隣接するサブフレーム間では同じような形状が繰り返されるため、これらのデータを残差信号の第1段目の近似に使

用することにより、より高精度な残差信号の符号化を行うことができるからである。なお、特には図示しない音声復号装置の側でも、全く同じようにしてアダプティブコードブックの更新が行われるため、データに矛盾が発生することはない。

【0106】以上の記載では、本発明を音声符号化方式の1方式であるCELP方式に適用した実施例について説明したが、本発明はこれに限られることなく、例えば残差信号の符号化方式の1つであるマルチパルス符号化方式において、A-b-S法に基づいて入力音声と合成された音声の二乗誤差の評価を繰り返す処理などについても、同様に適用することができる。

【0107】

【発明の効果】本発明によれば、音声信号は隣接する数サンプル間では近接相関が高いことを利用して、間引き手段により入力音声ベクトル及び合成音声ベクトルの各要素値の間引いて得られる間引き入力音声ベクトル及び間引き合成音声ベクトルについて誤差評価手段109で誤差評価が行われることにより、復号装置側での合成音声の品質の劣化を最小限に抑えつつ、誤差評価のための演算量を数分の1にすることが可能となり、音声符号化装置の高速化・回路規模の縮小が可能となる。

【図面の簡単な説明】

【図1】本発明のブロック図である。

【図2】本発明による音声符号化装置におけるLPC符号化部の構成図である。

【図3】本発明による音声符号化装置における残差符号化部の構成図である。

【図4】本発明による音声符号化装置における誤差評価部304の構成図である。

【図5】本発明による音声符号化装置における誤差評価部310の構成図である。

【図6】数12式と数13式の演算量を比較した図である。

【図7】音声信号と残差信号を示した図である。

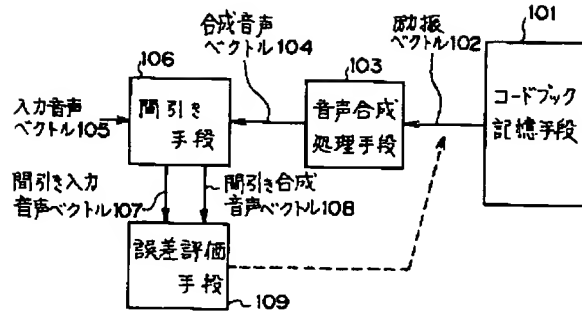
【図8】CELP方式の原理構成図である。

【符号の説明】

| | |
|-----|-------------|
| 101 | コードブック記憶手段 |
| 102 | 励振ベクトル |
| 103 | 音声合成処理手段 |
| 104 | 合成音声ベクトル |
| 105 | 入力音声ベクトル |
| 106 | 間引き手段 |
| 107 | 間引き入力音声ベクトル |
| 108 | 間引き合成音声ベクトル |
| 109 | 誤差評価手段 |

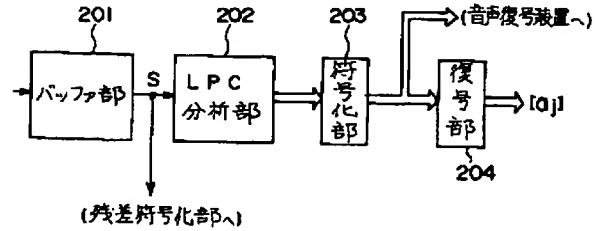
【図1】

本発明のブロック図



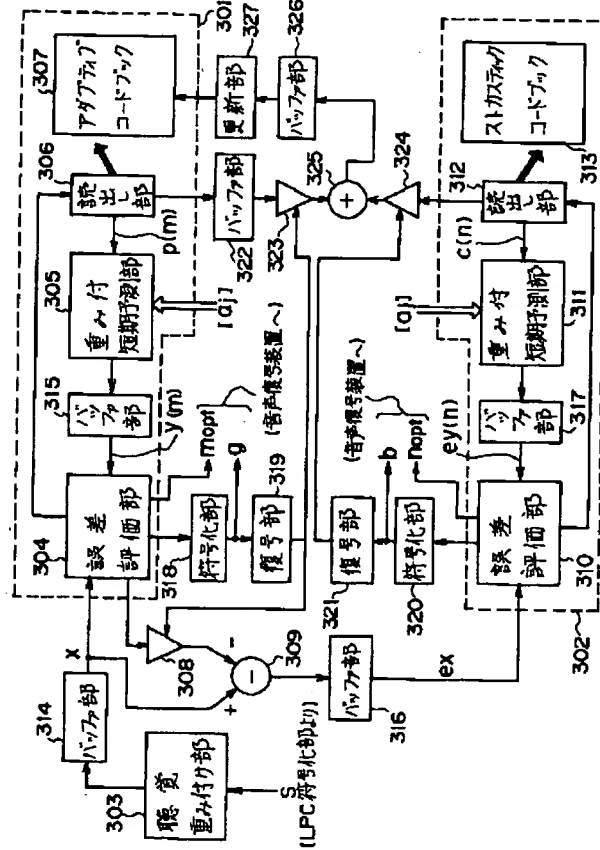
【図2】

本発明による音声符号化装置におけるLPC符号化部の構成図



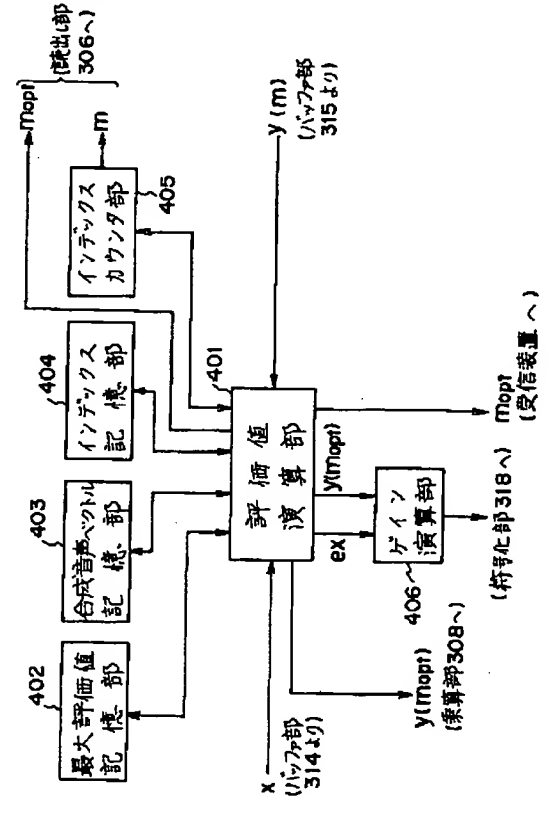
【図3】

本発明による音声符号化装置における残差符号化部の構成図



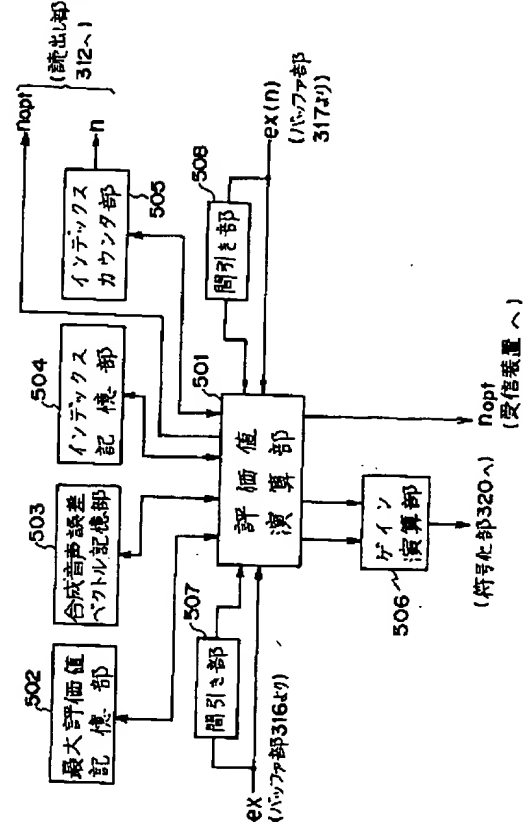
【図4】

本発明による音声符号化装置における誤差評価部304の構成図



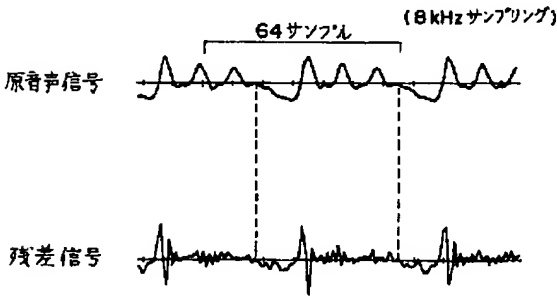
【図5】

本発明による音声符号化装置における誤差評価部310の構成図



【図7】

音声信号と残差信号を示した図



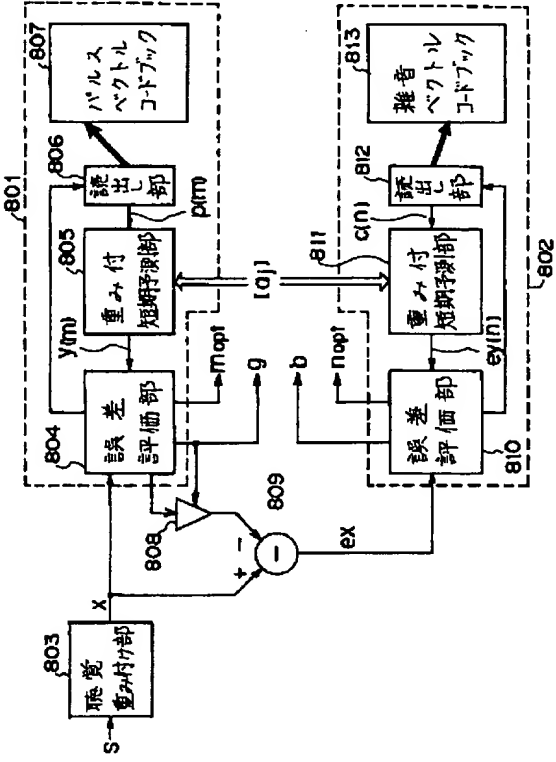
【図6】

数12式と数13式の演算量を比較した図

| | 数12式 | 数13式 |
|----------------|----------------------|----------------------------|
| 加算回数 | $(N-1)+(N-1)=2(N-1)$ | $(N/M-1)+(N/M-1)=2(N/M-1)$ |
| 乗算回数 (除算含む) | $(N+1)+N+1=2(N+1)$ | $(N/M+1)+N/M+1=2(N/M+1)$ |

【図8】

CELP方式の原理構成図



フロントページの続き

(72)発明者 栗原 秀明
神奈川県川崎市中原区上小田中1015番地
富士通株式会社内

(72)発明者 天野 文雄
神奈川県川崎市中原区上小田中1015番地
富士通株式会社内